GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 14:38:11 ; Search time 1752 Seconds
(without alignments)
811.122 Million cell updates/sec

Title: US-10-081-555C-3
Perfect score: 25
Sequence: 1 tagacagttcatgaagttcatctac 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	25	100.0	25	6	AR561609	AR561609 Sequence
	2	25	100.0	25	6	AR593234	AR593234 Sequence
	3	25	100.0	25	6	AX172812	AX172812 Sequence
c	4	25	100.0	360	9	S82239	S82239 CYP3A23=maj
c	5	25	100.0	1700	9	RATP450P	M86850 Rattus norv
c	6	25	100.0	2073	6	AX827858	AX827858 Sequence
c	7	25	100.0	2073	9	RNCYP3A1	X62086 R.norvegicu
c	8	25	100.0	7562	9	AB008388	AB008388 Rattus no
c	9	25	100.0	175876	14	AC123336	AC123336 Rattus no
c	10	25	100.0	263127	14	AC112327	AC112327 Rattus no
	11	23	92.0	31	6	BD225215	BD225215 Orphan nu
	12	23	92.0	31	6	AX399455	AX399455 Sequence
	13	23	92.0	32	6	BD225220	BD225220 Orphan nu
	14	23	92.0	32	6	AX399460	AX399460 Sequence
c	15	21.8	87.2	4230	9	RNCYTOBA01	U09725 Rattus norv
c	16	21.8	87.2	8014	9	RNTEST6	X79319 R.norvegicu
c	17	21	84.0	21	6	BD225225	BD225225 Orphan nu
c	18	21	84.0	21	6	BD227104	BD227104 Biologica
	19	20.4	81.6	60858	8	AL512592	AL512592 Human DNA
	20	20.4	81.6	174040	14	AC073950	AC073950 Homo sapi
	21	20.4	81.6	217652	14	AC150621	AC150621 Callithri
	22	20.2	80.8	25	6	AR561610	AR561610 Sequence
	23	20.2	80.8	25	6	AR593235	AR593235 Sequence
	24	20.2	80.8	25	6	AX172813	AX172813 Sequence
c	25	20.2	80.8	1572	9	RNCYP3A2	X62087 R.norvegicu
	26	19.8	79.2	110000	1	AE017261_10	Continuation (11 o
	27	19.8	79.2	170588	8	CNS01RHF	AL161747 Human chr
c	28	19.8	79.2	184696	8	AC112507	AC112507 Homo sapi
c	29	19.8	79.2	216402	5	AC145978	AC145978 Gallus ga
	30	19.8	79.2	233033	5	AC145947	AC145947 Gallus ga
c	31	19.4	77.6	986	5	BX934150	BX934150 Gallus ga
	32	19.4	77.6	66374	15	CR382128_30	Continuation (31 o
c	33	19.2	76.8	2079	2	AF276833	AF276833 Phytomyza
c	34	19.2	76.8	71111	14	CR457451_3	Continuation (4 of
c	35	19.2	76.8	110000	1	AE016853_48	Continuation (49 o
c	36	19.2	76.8	110000	14	TANN2_02	Continuation (3 of
c	37	19.2	76.8	125205	9	AC154022	AC154022 Mus muscu
	38	19.2	76.8	138943	14	CR847865	CR847865 Danio rer
c	39	19.2	76.8	159798	5	BX323794	BX323794 Zebrafish
c	40	19.2	76.8	163765	5	BX284684	BX284684 Zebrafish
	41	19.2	76.8	174551	8	AC010598	AC010598 Homo sapi
	42	19.2	76.8	200139	14	CR376839	CR376839 Danio rer
	43	19.2	76.8	201349	14	CR847975	CR847975 Danio rer
c	44	19.2	76.8	218884	14	AC161536	AC161536 Mus muscu
	45	19.2	76.8	227533	9	AC102656	AC102656 Mus muscu

ALIGNMENTS

RESULT 1
AR561609
LOCUS AR561609 25 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6756491.
ACCESSION AR561609
VERSION AR561609.1 GI:53974716
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 25)
AUTHORS Evans,R.M. and Blumberg,B.
TITLE Steroid-activated nuclear receptors and uses therefor
JOURNAL Patent: US 6756491-A 3 29-JUN-2004;
The Salk Institute for Biological Studies; La Jolla, CA
FEATURES Location/Qualifiers
source 1. .25
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAGACAGTTCATGAAGTTCATCTAC 25
|||||||||||||||||||||||||
Db 1 TAGACAGTTCATGAAGTTCATCTAC 25

RESULT 2
AR593234
LOCUS AR593234 25 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 3 from patent US 6809178.
ACCESSION AR593234
VERSION AR593234.1 GI:56642319
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 25)

```
  AUTHORS   Evans,R.M. and Blumberg,B.
  TITLE     Steroid-activated nuclear receptors and uses therefor
  JOURNAL   Patent: US 6809178-A 3 26-OCT-2004;
            The Salk Institute for Biological Studies; La Jolla, CA
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           100.0%;  Score 25;  DB 6;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 2.6;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db          1 TAGACAGTTCATGAAGTTCATCTAC 25


RESULT 3
AX172812
LOCUS       AX172812                  25 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 3 from Patent WO0142290.
ACCESSION   AX172812
VERSION     AX172812.1  GI:14597861
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Evans,R.M., Blumberg,B. and Xie,W.
  TITLE     Novel steroid-activated nuclear receptors and uses therefor
  JOURNAL   Patent: WO 0142290-A 3 14-JUN-2001;
            THE SALK INSTITUTE FOR BIOLOGICAL STUDIES (US)
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="putative SXR response element from the steroid
                     hydroxylase, rCYP3A1"
ORIGIN

  Query Match           100.0%;  Score 25;  DB 6;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 2.6;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db          1 TAGACAGTTCATGAAGTTCATCTAC 25


RESULT 4
S82239/c
LOCUS       S82239                   360 bp    DNA     linear   ROD 12-FEB-1997
DEFINITION  CYP3A23=major glucocorticoid-inducible cytochrome P3A {promoter}
            [rats, Wistar-Furth, Genomic, 360 nt].
ACCESSION   S82239
VERSION     S82239.1  GI:1839503
KEYWORDS    .
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 360)
  AUTHORS   Huss,J.M., Wang,S.I., Astrom,A., McQuiddy,P. and Kasper,C.B.
  TITLE     Dexamethasone responsiveness of a major glucocorticoid-inducible
            CYP3A gene is mediated by elements unrelated to a glucocorticoid
            receptor binding motif
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (10), 4666-4670 (1996)
  PUBMED    8643461
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 178156] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .360
                     /organism="Rattus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10118"
     gene            1. .360
                     /gene="CYP3A23"
                     /note="major glucocorticoid-inducible cytochrome P3A"
     mRNA            304. .>360
                     /gene="CYP3A23"
ORIGIN

  Query Match           100.0%;  Score 25;  DB 9;  Length 360;
  Best Local Similarity  100.0%;  Pred. No. 2.1;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db        189 TAGACAGTTCATGAAGTTCATCTAC 165


RESULT 5
RATP450P/c
LOCUS       RATP450P                 1700 bp    DNA     linear   ROD 30-MAY-2000
DEFINITION  Rattus norvegicus cytochrome P-450 (CYP3A1) gene, partial cds.
ACCESSION   M86850
VERSION     M86850.1  GI:205919
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 1700)
  AUTHORS   Burger,H.J., Schuetz,J.D., Schuetz,E.G. and Guzelian,P.S.
  TITLE     Paradoxical transcriptional activation of rat liver cytochrome
            P-450 3A1 by dexamethasone and the antiglucocorticoid pregnenolone
            16 alpha-carbonitrile: analysis by transient transfection into
            primary monolayer cultures of adult rat hepatocytes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (6), 2145-2149 (1992)
  PUBMED    1372436
FEATURES             Location/Qualifiers
     source          1. .1700
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /tissue_lib="EMBL1-1-4"
     gene            1495. .1700
                     /gene="CYP3A1"
     TATA_signal     1495. .1499
                     /gene="CYP3A1"
     CDS             1616. .>1686
                     /gene="CYP3A1"
                     /codon_start=1
                     /product="cytochrome P-450"
                     /protein_id="AAA41780.1"
                     /db_xref="GI:205920"
                     /translation="MDLLSALTLETWVLLAVVLVLLYG"
ORIGIN

  Query Match           100.0%;  Score 25;  DB 9;  Length 1700;
  Best Local Similarity  100.0%;  Pred. No. 1.9;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db       1411 TAGACAGTTCATGAAGTTCATCTAC 1387


RESULT 6
AX827858/c
LOCUS       AX827858                 2073 bp    DNA     linear   PAT 12-DEC-2003
```

```
DEFINITION  Sequence 592 from Patent EP1344834.
ACCESSION   AX827858
VERSION     AX827858.1  GI:39838046
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Boess,F., Suter-Dick,L. and Wolf,D.
  TITLE     Methods for the toxicity prediction of a compound
  JOURNAL   Patent: EP 1344834-A 592 17-SEP-2003;
            F. HOFFMANN-LA ROCHE AG (CH)
FEATURES             Location/Qualifiers
     source          1. .2073
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             100.0%;  Score 25;  DB 6;  Length 2073;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db       1284 TAGACAGTTCATGAAGTTCATCTAC 1260


RESULT 7
RNCYP3A1/c
LOCUS       RNCYP3A1                2073 bp    DNA     linear   ROD 10-AUG-1992
DEFINITION  R.norvegicus CYP3A1 gene for cytochrome P450 PCN1.
ACCESSION   X62086
VERSION     X62086.1  GI:56037
KEYWORDS    CYP3A1 gene; Cytochrome P450 PCN1; dexamethosone-induced cytochrome
            P450; monoxygenase; NADP dependent cytochrome P450 reductase;
            phenobarbital-induced cytochrome P-450.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Telhada,M.B., Pereira,T.M. and Lechner,M.C.
  TITLE     Effect of dexamethasone and phenobarbital on run-on transcription
            rate and CYP3A mRNA concentration in rat liver: changes during
            development
  JOURNAL   Arch. Biochem. Biophys.
REFERENCE   2  (bases 1 to 2073)
  AUTHORS   Lechner,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1991) M.C. Lechner, Instituto Gulbenkian de
            Ciencia, Lab. Bioquimica, Apartado 14, 2781 Oeiras Codex, PORTUGAL
COMMENT     For related sequences see X62087, M10161, Gonzalez F.J.; Mol.Cell
            Biol 6:2969-2976 (1986) & Yanagida A.; Mol.Cell Biol.
            10:1470-1475(1990).
FEATURES             Location/Qualifiers
     source          1. .2073
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /clone="pBluescript II SK +/- CYP3A1"
                     /tissue_type="liver"
                     /clone_lib="genomic: EMBL3cosW"
                     /dev_stage="adult"
     misc_feature    complement(551. .562)
                     /note="Basic transcription element"
     TATA_signal     1367. .1371
     gene            1398. .1557
                     /gene="CYP3A1"
     mRNA            1398. .1557
                     /gene="CYP3A1"
                     /note="product EC number is 1.14.14.1"
     exon            1398. .1557
                     /gene="CYP3A1"
                     /number=1
ORIGIN

  Query Match             100.0%;  Score 25;  DB 9;  Length 2073;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db       1284 TAGACAGTTCATGAAGTTCATCTAC 1260


RESULT 8
AB008388/c
LOCUS       AB008388                7562 bp    DNA     linear   ROD 27-OCT-2004
DEFINITION  Rattus norvegicus CYP3A1 gene for cytochrome P450/6 beta B,
            complete cds.
ACCESSION   AB008388 AB008377 AB008378 AB008379 AB008380 AB008381 AB008382
            AB008383 AB008384 AB008385 AB008386 AB008387 AB008389
VERSION     AB008388.2  GI:60391376
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Nagata,K., Ogino,M., Shimada,M., Miyata,M., Gonzalez,F.J. and
            Yamazoe,Y.
  TITLE     Structure and expression of the rat CYP3A1 gene: isolation of the
            gene (P450/6betaB) and characterization of the recombinant protein
  JOURNAL   Arch. Biochem. Biophys. 362 (2), 242-253 (1999)
   PUBMED   9989933
REFERENCE   2  (bases 1 to 7562)
  AUTHORS   Nagata,K., Ogino,M., Shimada,M., Miyata,M. and Yamazoe,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1997) Kiyoshi Nagata, Faculty of Pharmaceutical
            Sciences, Tohoku University, Division of Drug Metabolism and
            Molecular Toxicology; Aza-Aoba, Aramaki, Aoba-ku, Sendai, Miyagi,
            980-77, Japan (E-mail:nagataki@mail.pharm.tohoku.ac.jp,
            Tel:022-217-6830, Fax:022-217-6826)
COMMENT     On or before Mar 1, 2005 this sequence version replaced gi:2575801,
            gi:2575789, gi:2575790, gi:2575791, gi:2575792, gi:2575793,
            gi:2575794, gi:2575795, gi:2575796, gi:2575797, gi:2575798,
            gi:2575799, gi:2575800.
FEATURES             Location/Qualifiers
     source          1. .7562
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
                     /dev_stage="adult"
     TATA_signal     1093. .1097
     gap             1123. .1222
                     /estimated_length=unknown
     gene            1223. .7519
                     /gene="CYP3A1"
     exon            1223. .1383
                     /gene="CYP3A1"
                     /number=1
     CDS             join(1313. .1383,1795. .1888,2295. .2347,2668. .2767,
                     3266. .3379,3457. .3539,3910. .4058,4920. .5047,5349. .5418,
                     5746. .5906,6110. .6336,6494. .6656,6983. .7078)
                     /gene="CYP3A1"
                     /codon_start=1
                     /product="cytochrome P450/6 beta B"
                     /protein_id="BAA23003.1"
```

```
                     /db_xref="GI:2575802"
                     /translation="MDLLSALTLETWVLLAVVLVLLYGFGTRTHGLFKKQGIPGPKPL
                     PFFGTVLNYYMGLWKFDVECHKKYGKIWGLFDGQMPLFAITDTEMIKNVLVKECFSVF
                     TNRRDFGPVGIMGKAISVSKDEEWKRYRALLSPTFTSGRLKEMFPVIEQYGDILVKYL
                     RQEKGKPVPVKEVFGAYSMDVITSTSFGVNVDSLNNPKDPFVEKAKKLLRIDFFDPLF
                     LSVVLFPFLTPVYEMLNICMFPKDSIEFFKKFVYRMKETRLDSVQKHRVDFLQLMMNA
                     HNDSKDKESHTALSDMEITAQSIIFIFAGYEPTSSTLSFVLHSLATHPDTQKKLQEEI
                     DRALPNKAPPTYDTVMEMEYLDMVLNETLRLYPIGNRLERVCKKDVEINGVFMPKGSV
                     VMIPSYALHRDPQHWPEPEEFRPERFSKENKGSIDPYVYLPFGNGPRNCIGMRFALMN
                     MKLALTKVLQNFSFQPCKETQIPLKLSRQGLLQPTKPIILKVVPRDEIITGS"
     gap             1620. .1719
                     /estimated_length=unknown
     exon            1795. .1888
                     /gene="CYP3A1"
                     /number=2
     gap             2013. .2112
                     /estimated_length=unknown
     exon            2295. .2347
                     /gene="CYP3A1"
                     /number=3
     gap             2478. .2577
                     /estimated_length=unknown
     exon            2668. .2767
                     /gene="CYP3A1"
                     /number=4
     gap             2881. .2980
                     /estimated_length=unknown
     exon            3266. .3379
                     /gene="CYP3A1"
                     /number=5
     exon            3457. .3539
                     /gene="CYP3A1"
                     /number=6
     gap             3654. .3753
                     /estimated_length=unknown
     exon            3910. .4058
                     /gene="CYP3A1"
                     /number=7
     gap             4645. .4744
                     /estimated_length=unknown
     exon            4920. .5047
                     /gene="CYP3A1"
                     /number=8
     gap             5062. .5161
                     /estimated_length=unknown
     exon            5349. .5418
                     /gene="CYP3A1"
                     /number=9
     gap             5580. .5679
                     /estimated_length=unknown
     exon            5746. .5906
                     /gene="CYP3A1"
                     /number=10
     gap             5981. .6080
                     /estimated_length=unknown
     exon            6110. .6336
                     /gene="CYP3A1"
                     /number=11
     gap             6346. .6445
                     /estimated_length=unknown
     exon            6494. .6656
                     /gene="CYP3A1"
                     /number=12
     gap             6712. .6811
                     /estimated_length=unknown
     exon            6983. .7519
                     /gene="CYP3A1"
                     /number=13
ORIGIN

  Query Match           100.0%;  Score 25;  DB 9;  Length 7562;
  Best Local Similarity 100.0%;  Pred. No. 1.7;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db       1009 TAGACAGTTCATGAAGTTCATCTAC 985


RESULT 9
AC123336/c
LOCUS       AC123336              175876 bp    DNA     linear   HTG 20-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-264B4, WORKING DRAFT SEQUENCE, 3
            unordered pieces.
ACCESSION   AC123336
VERSION     AC123336.4  GI:25138142
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 175876)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175876)
  AUTHORS   Worley,K.C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 175876)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:23811860.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GUND
                Center clone name: CH230-264B4
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 163335 bases at least Q40
                Consensus quality: 165141 bases at least Q30
                Consensus quality: 166401 bases at least Q20
                Estimated insert size: 169316; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    25872: contig of 25872 bp in length
            *    25873    25972: gap of unknown length
            *    25973   174283: contig of 148311 bp in length
            *   174284   174383: gap of unknown length
            *   174384   175876: contig of 1493 bp in length.
FEATURES             Location/Qualifiers
     source          1. .175876
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-264B4"
     misc_feature    2416. .3273
                     /note="clone_boundary
                     clone_end:Sp6
                     site:
                     end_sequence:RXADX02TV"
     gap             25873. .25972
                     /estimated_length=unknown
     misc_feature    25973. .27022
                     /note="wgs_contig"
     misc_feature    complement(142671. .143326)
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:BZ120347"
     misc_feature    169909. .171359
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    173203. .174283
                     /note="wgs_end_extension
                     clone_end:T7"
     gap             174284. .174383
                     /estimated_length=unknown
ORIGIN

  Query Match             100.0%;  Score 25;  DB 14;  Length 175876;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db     115324 TAGACAGTTCATGAAGTTCATCTAC 115300


RESULT 10
AC112327/c
LOCUS       AC112327               263127 bp    DNA     linear   HTG 09-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-177H19, *** SEQUENCING IN PROGRESS
            ***.
ACCESSION   AC112327
VERSION     AC112327.4  GI:24635585
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 263127)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
```

```
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 263127)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 263127)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 6, 2002 this sequence version replaced gi:23101322.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GOVD
                Center clone name: CH230-177H19
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 203179 bases at least Q40
                Consensus quality: 206343 bases at least Q30
                Consensus quality: 208281 bases at least Q20
                Estimated insert size: 208801; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   263127: contig of 263127 bp in length.
FEATURES            Location/Qualifiers
     source         1. .263127
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-177H19"
     misc_feature   1. .1099
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   30793. .32104
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   32283. .32865
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH273139"
ORIGIN

  Query Match            100.0%;  Score 25;  DB 14;  Length 263127;
  Best Local Similarity  100.0%;  Pred. No. 1.3;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db      62954 TAGACAGTTCATGAAGTTCATCTAC 62930


RESULT 11
BD225215
LOCUS       BD225215                  31 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Orphan nuclear receptor.
ACCESSION   BD225215
VERSION     BD225215.1  GI:33034985
KEYWORDS    JP 2002535241-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Kliewer,S.A. and Willson,T.M.
  TITLE     Orphan nuclear receptor
  JOURNAL   Patent: JP 2002535241-A 4 22-OCT-2002;
            GLAXO GROUP LTD
COMMENT     OS   Artificial Sequence
            PN   JP 2002535241-A/4
            PD   22-OCT-2002
            PF   26-MAR-1999 JP 2000537897
            PR   27-MAR-1998 US    60/079593
            PI   STEVEN ANTHONY KLIEWER,TIMOTHY MARK WILLSON
            PC   C07K14/00,C07K14/435,C07K14/705,C07K19/00,C12N5/10,C12N15/09,
            PC   C12Q1/68,
            PC   G01N33/15,G01N33/50,G01N33/566,C12N5/00,C12N15/00 CC   DNA
            genome
            FH   Key             Location/Qualifiers
            FT   source          1. .31
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .31
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
ORIGIN

  Query Match            92.0%;  Score 23;  DB 6;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 21;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTA 24
              |||||||||||||||||||||||
Db          5 AGACAGTTCATGAAGTTCATCTA 27


RESULT 12
AX399455
LOCUS       AX399455                  31 bp    DNA     linear   PAT 28-MAY-2002
DEFINITION  Sequence 4 from Patent WO0197856.
ACCESSION   AX399455
VERSION     AX399455.1  GI:21262007
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
```

```
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Kliewer,S.A., Jones,S.A. and Willson,T.M.
  TITLE     An orphan nuclear receptor
  JOURNAL   Patent: WO 0197856-A 4 27-DEC-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="DNA genome"
ORIGIN

  Query Match           92.0%;  Score 23;  DB 6;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 21;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTA 24
              |||||||||||||||||||||||
Db          5 AGACAGTTCATGAAGTTCATCTA 27


RESULT 13
BD225220
LOCUS       BD225220                  32 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Orphan nuclear receptor.
ACCESSION   BD225220
VERSION     BD225220.1  GI:33034990
KEYWORDS    JP 2002535241-A/9.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Kliewer,S.A. and Willson,T.M.
  TITLE     Orphan nuclear receptor
  JOURNAL   Patent: JP 2002535241-A 9 22-OCT-2002;
            GLAXO GROUP LTD
COMMENT     OS   Artificial Sequence
            PN   JP 2002535241-A/9
            PD   22-OCT-2002
            PF   26-MAR-1999 JP 2000537897
            PR   27-MAR-1998 US    60/079593
            PI   STEVEN ANTHONY KLIEWER,TIMOTHY MARK WILLSON
            PC   C07K14/00,C07K14/435,C07K14/705,C07K19/00,C12N5/10,C12N15/09,
            PC   C12Q1/68,
            PC   G01N33/15,G01N33/50,G01N33/566,C12N5/00,C12N15/00 CC   DNA
            genome
            FH   Key             Location/Qualifiers
            FT   source          1. .32
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           92.0%;  Score 23;  DB 6;  Length 32;
  Best Local Similarity  100.0%;  Pred. No. 21;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTA 24
              |||||||||||||||||||||||
Db          6 AGACAGTTCATGAAGTTCATCTA 28


RESULT 14
AX399460
LOCUS       AX399460                  32 bp    DNA     linear   PAT 28-MAY-2002
DEFINITION  Sequence 9 from Patent WO0197856.
ACCESSION   AX399460
VERSION     AX399460.1  GI:21262012
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Kliewer,S.A., Jones,S.A. and Willson,T.M.
  TITLE     An orphan nuclear receptor
  JOURNAL   Patent: WO 0197856-A 9 27-DEC-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="DNA genome"
ORIGIN

  Query Match           92.0%;  Score 23;  DB 6;  Length 32;
  Best Local Similarity  100.0%;  Pred. No. 21;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTA 24
              |||||||||||||||||||||||
Db          6 AGACAGTTCATGAAGTTCATCTA 28


RESULT 15
RNCYTOBA01/c
LOCUS       RNCYTOBA01              4230 bp    DNA     linear   ROD 22-JUN-1995
DEFINITION  Rattus norvegicus testosterone 6-beta-hydroxylase, cytochrome
            P450/6-beta-A, (CYP3A2) gene, exons 1 and 2.
ACCESSION   U09725 M74443
VERSION     U09725.1  GI:498847
KEYWORDS    .
SEGMENT     1 of 10
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 4230)
  AUTHORS   Miyata,M., Nagata,K., Shimada,M., Yamazoe,Y. and Kato,R.
  TITLE     Structure of a gene and cDNA of a major constitutive form of
            testosterone 6 beta-hydroxylase (P450/6 beta A) encoding CYP3A2:
            comparison of the cDNA with P450PCN2
  JOURNAL   Arch. Biochem. Biophys. 314 (2), 351-359 (1994)
   PUBMED   7979376
REFERENCE   2  (bases 3311 to 3569)
  AUTHORS   Miyata,M., Nagata,K., Yamazoe,Y. and Kato,R.
  TITLE     A gene structure of testosterone 6 beta-hydroxylase (P450IIIA)
  JOURNAL   Biochem. Biophys. Res. Commun. 177 (1), 68-73 (1991)
   PUBMED   2043144
REFERENCE   3  (bases 1 to 4230)
  AUTHORS   Miyata,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-1994) Masaaki Miyata, Department of Pharmacology,
            Keio University, School of Medicine, 35 Shinanomachi, Shinjuku-ku,
            Tokyo 160, Japan
COMMENT     On Jun 13, 1994 this sequence version replaced gi:205918.
FEATURES             Location/Qualifiers
     source          1. .4230
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
                     /dev_stage="adult"
     exon        .   3392. .3530
                     /gene="CYP3A2"
                     /note="residues 3392. .3459 are not translated"
                     /number=1
     conflict        3555^3556
                     /gene="CYP3A2"
```

```
                     /note="sequencing error in GenBank Accession Number
                     M74443"
                     /citation=[2]
                     /replace="T"
     exon            3900. .3993
                     /gene="CYP3A2"
                     /number=2
ORIGIN

  Query Match             87.2%;  Score 21.8;  DB 9;  Length 4230;
  Best Local Similarity   92.0%;  Pred. No. 51;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              || ||||||||| ||||||||||||
Db       3256 TAAACAGTTCATAAAGTTCATCTAC 3232


Search completed: February  6, 2006, 15:11:32
Job time : 1756 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 14:20:25 ; Search time 286 Seconds
(without alignments)
582.578 Million cell updates/sec

Title: US-10-081-555C-3
Perfect score: 25
Sequence: 1 tagacagttcatgaagttcatctac 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	25	100.0	25	2	AAX89081	Aax89081 Putative
	2	25	100.0	25	5	AAH25490	Aah25490 Steroid-a
	3	25	100.0	25	8	ABZ58304	Abz58304 Direct re
	4	25	100.0	25	9	ACD27769	Acd27769 Steroid h
	5	25	100.0	25	9	ACD40529	Acd40529 Rat stero
	6	25	100.0	25	10	AAD50114	Aad50114 Rat CYP3A
c	7	25	100.0	2073	11	ADW22213	Adw22213 Rat hepat
	8	23	92.0	31	2	AAZ07991	Aaz07991 Oligo con
	9	23	92.0	31	6	ABA91215	Aba91215 CYP3A1 DR
	10	23	92.0	32	2	AAZ07996	Aaz07996 Radiolabe
c	11	21	84.0	21	3	AAZ40699	Aaz40699 Rat CYP3A
	12	20.2	80.8	25	2	AAX89082	Aax89082 Putative
	13	20.2	80.8	25	5	AAH25491	Aah25491 Steroid-a
	14	20.2	80.8	25	8	ABZ58305	Abz58305 Direct re
	15	20.2	80.8	25	9	ACD27770	Acd27770 Steroid h
	16	20.2	80.8	25	9	ACD40530	Acd40530 Rat stero
	17	20.2	80.8	25	10	AAD50115	Aad50115 Rat CYP3A
	18	18.6	74.4	2000	8	ADA71815	Ada71815 Rice gene
	19	18.6	74.4	2000	11	ACL35429	Acl35429 Rice stre
c	20	18.6	74.4	4479	12	ADH22292	Adh22292 Rice PONG
	21	18.4	73.6	778	4	AAI95094	Aai95094 Human neu
	22	18.4	73.6	778	8	ABT42837	Abt42837 Human neu
	23	18.2	72.8	6363	4	AAK76485	Aak76485 Human imm
	24	18.2	72.8	96596	9	ADA02504	Ada02504 Human BAC
	25	18.2	72.8	96596	10	ADB72242	Adb72242 Human BAC
	26	18.2	72.8	96596	10	ADE95752	Ade95752 Human BAC
	27	18.2	72.8	160482	11	ACN43914	Acn43914 Human gen
c	28	17.8	71.2	681	3	AAF13979	Aaf13979 Aspergill
c	29	17.8	71.2	681	13	ADU58020	Adu58020 Aspergill
c	30	17.8	71.2	681	14	ADZ96023	Adz96023 Aspergill
c	31	17.8	71.2	700	6	ABK62828	Abk62828 Rat seque
c	32	17.8	71.2	700	10	ADB51335	Adb51335 Primary r
c	33	17.8	71.2	700	10	ABT41183	Abt41183 Toxicity
c	34	17.8	71.2	700	12	ADP72122	Adp72122 Renal tox
	35	17.8	71.2	990	13	ADT45639	Adt45639 Bacterial
	36	17.8	71.2	993	13	ADS46582	Ads46582 Bacterial
c	37	17.8	71.2	1579	11	ADM03304	Adm03304 Human cDN
	38	17.6	70.4	400	5	ABV53868	Abv53868 Human pro
	39	17.6	70.4	551	13	ADX64878	Adx64878 Plant ful
	40	17.6	70.4	651	3	AAF11281	Aaf11281 Aspergill
	41	17.6	70.4	651	13	ADU55322	Adu55322 Aspergill
	42	17.6	70.4	651	14	ADZ93325	Adz93325 Aspergill
	43	17.6	70.4	825	10	ADC91683	Adc91683 E. faeciu
c	44	17.6	70.4	2371	3	AAA35186	Aaa35186 Human ade
c	45	17.6	70.4	2371	3	AAF21308	Aaf21308 Human low

ALIGNMENTS

RESULT 1
AAX89081
ID AAX89081 standard; DNA; 25 BP.
XX
AC AAX89081;
XX
DT 14-SEP-1999 (first entry)
XX
DE Putative SXR response element DR-3 containing fragment rCYP3A1.
XX
KW Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;
KW retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;
KW phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;
KW breast cancer; osteoporosis; Cushing syndrome; virilism; hirsutism;
KW polycystic ovarian disease; cancer; colorectal; prostatic; ss.
XX
OS Homo sapiens.
XX
PN WO9935246-A1.
XX
PD 15-JUL-1999.
XX
PF 08-JAN-1999; 99WO-US000490.
XX
PR 09-JAN-1998; 98US-00005286.
XX
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Evans RM, Blumberg B;
XX
DR WPI; 1999-419349/35.
XX
PT New steroid and xenobiotic receptor, used to identify modulators for
PT controlling metabolism of steroids and xenobiotics, e.g. reducing their
PT toxicity.
XX
PS Disclosure; Fig 6A; 83pp; English.
XX
CC The invention relates to a novel nuclear receptor polypeptide, designated
CC SXR (steroid and xenobiotic receptor). SXR (i) forms a heterodimer with
CC retinoid X receptor (RXR), (ii) binds to a direct or inverted repeat
CC response element motif based on the half-site AGTTCA, (iii) activates

```
CC   transcription through response elements present in steroid-inducible P450
CC   genes, in response to a wide variety of natural and synthetic steroid
CC   hormones and (iv) is prominently expressed in liver and intestine. SXR
CC   regulates expression of catabolic enzymes, in response to many different
CC   steroids, and thus affects metabolism. SXR is a broad specificity, low-
CC   affinity receptor for reducing excessive levels of steroids in the
CC   circulation (see AAX89090 for detailed uses of SXR polypeptide).
CC   Sequences AAX89081-89 represent fragments from various steroid and
CC   xenobiotic inducible P450 enzymes containing putative SXR response
CC   elements
XX
SQ   Sequence 25 BP; 8 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 25;  DB 2;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db          1 TAGACAGTTCATGAAGTTCATCTAC 25


RESULT 2
AAH25490
ID   AAH25490 standard; DNA; 25 BP.
XX
AC   AAH25490;
XX
DT   22-AUG-2001  (first entry)
XX
DE   Steroid-activated nuclear receptor putative response element.
XX
KW   Steroid-activated nuclear receptor; steroid and xenobiotic receptor; SXR;
KW   retinoid X receptor; RXR; transcription; response element;
KW   steroid inducible P450 gene; steroid hormone; Cushing's syndrome;
KW   virulism; hirsutism; polycystic ovarian syndrome; hypertension; ss.
XX
OS   Unidentified.
XX
PN   WO200142290-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033473.
XX
PR   09-DEC-1999;   99US-00458366.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Evans RM,  Blumberg B,  Xie W;
XX
DR   WPI; 2001-381637/40.
XX
PT   Novel steroid-activated nuclear receptor useful as sensor for xenobiotic
PT   compounds and/or steroids and whose modulators are useful for modulating
PT   metabolism of steroid or xenobiotic compounds.
XX
PS   Disclosure; Page 23; 64pp; English.
XX
CC   The present sequence represents a putative response element for a steroid
CC   -activated nuclear receptor, termed steroid and xenobiotic receptor
CC   (SXR). The response element is identified in steroid hydroxylase CYP3A1.
CC   The SXR polypeptide is capable of forming a heterodimer with retinoid X
CC   receptor (RXR), activating transcription through response elements found
CC   in steroid inducible P450 genes in response to a variety of natural and
CC   synthetic steroid hormones and prominently expressed in liver and
CC   intestine. SXR binds to a direct or inverted repeat response element
CC   motif based on the half site AGTTCA. SXR is useful for identifying
CC   compounds which are agonists or which activate the receptor. The
CC   compounds identified are useful for treating a wide variety of conditions
CC   such as Cushing's syndrome, virulism and hirsutism, androgen excess due
CC   to polycystic ovarian syndrome and enzymatic defects which leads to
CC   accumulation of steroids, resulting in hypertension and aberrant
CC   development of secondary sexual characteristics in both sexes. Transgenic
CC   animals which express human SXR serve as models for human response to
CC   various agents which potentially impact P450-dependent metabolic
CC   processes
XX
SQ   Sequence 25 BP; 8 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 25;  DB 5;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db          1 TAGACAGTTCATGAAGTTCATCTAC 25


RESULT 3
ABZ58304
ID   ABZ58304 standard; DNA; 25 BP.
XX
AC   ABZ58304;
XX
DT   28-APR-2003  (first entry)
XX
DE   Direct repeat-3 in rat cytochrome P450-3A1 gene.
XX
KW   Steroid xenobiotic receptor; SXR; receptor; cytochrome-P450; rat;
KW   steroid; xenobiotic; antidote; detoxification; ds.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   repeat_region    6. .20
FT                    /*tag=  a
FT                    /rpt_type= DIRECT
FT                    /function= "Response element"
FT   repeat_unit      6. .11
FT                    /*tag=  b
FT   repeat_unit      15. .20
FT                    /*tag=  c
XX
PN   WO2003005812-A2.
XX
PD   23-JAN-2003.
XX
PF   09-JUL-2002; 2002WO-US021800.
XX
PR   09-JUL-2001; 2001US-0304388P.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Evans R,  Xie W;
XX
DR   WPI; 2003-221630/21.
XX
PT   Modulating the metabolism of steroids and xenobiotics with a UGT
PT   modulator, useful for modifying the physiological response to and/or
PT   efficient detoxification of harmful steroids and/or xenobiotic compounds.
XX
PS   Disclosure; Page 26; 51pp; English.
XX
CC   The present sequence is a direct repeat-3 (DR-3) response element for the
CC   steroid xenobiotic receptor (SXR/PXR) identified in the rat cytochrome
CC   P450-3A1 gene. A database search showed that putative SXR response
CC   elements are found in genes encoding steroid hydroxylases, P450
CC   oxidoreductase, and glucuronosyl transferase. SXR is a broad specificity,
CC   low affinity, steroid-activated receptor. The present invention relates
CC   to modulation of metabolism of steroids and xenobiotics. Nuclear
CC   receptors including SXR and constitutively active receptor (CAR) are
CC   characterised as xenosensors regulating expression of P450 genes. The
CC   ability of this group of receptors to regulate expression of UDP-
CC   glucuronosyl transferase (UGT) in response to steroids and/or xenobiotics
CC   provides novel approaches for direct regulation/activation of a
```

CC glucuronidation pathway, thereby providing methods to achieve physiologic
CC homeostasis with respect to steroids and/or xenobiotics. SXR and CAR
CC regulation of UGT represents the first evidence of receptors that can
CC transduce/transactivate both Phase I and Phase II adaptive hepatic
CC response. A claimed method for modulating the metabolism or clearance of
CC steroid and/or xenobiotic compounds involves administering a modulator of
CC UGT. The modulator can be a nucleic acid, protein and/or chemical
CC compound which binds a UGT direct repeat or inverted repeat response
CC element, or which activates UGT glucuronidation of steroids or
CC xenobiotics. A claimed method for inducing expression of steroid
CC hydroxylase comprises activating SXR/PXR and/or CAR
XX
SQ Sequence 25 BP; 8 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAGACAGTTCATGAAGTTCATCTAC 25
||||||||||||||||||||||||||
Db 1 TAGACAGTTCATGAAGTTCATCTAC 25

RESULT 4
ACD27769
ID ACD27769 standard; DNA; 25 BP.
XX
AC ACD27769;
XX
DT 18-SEP-2003 (first entry)
XX
DE Steroid hydroxylase rCYP3A1, putative SXR response element.
XX
KW Human; steroid X receptor; SXR; retinoid X receptor; RXR;
KW steroid inducible P450 gene; Cushing's syndrome; obesity; fatigue;
KW hypertension; oedema; osteoporosis; virilism; hirsutism; androgen excess;
KW polycystic ovarian syndrome; steroids accumulation; endocrine disruptor;
KW steroid hydroxylase; rCYP3A1; SXR response element; ds.
XX
OS Unidentified.
XX
PN US2003064430-A1.
XX
PD 03-APR-2003.
XX
PF 09-JAN-1998; 98US-00005286.
XX
PR 09-JAN-1998; 98US-00005286.
XX
PA (EVAN/) EVANS R M.
PA (BLUM/) BLUMBERG B.
XX
PI Evans RM, Blumberg B;
XX
DR WPI; 2003-540786/51.
XX
PT New steroid-activated nuclear receptor polypeptide that heterodimerizes
PT with retinoid X receptor, useful for identifying therapeutic compounds
PT for the treatment of Cushing's syndrome, virilism and hirsutism, and
PT androgen excess.
XX
PS Disclosure; Page 4; 23pp; English.
XX
CC The invention describes a new receptor polypeptide (I) or its functional
CC fragments. The method comprises forming a heterodimer with retinoid X
CC receptor (RXR), binding to a direct or inverted repeat response element
CC motif based on the half site AGTTCA, activating transcription through
CC response elements found in steroid inducible P450 genes in response to a
CC wide variety of natural and synthetic steroid hormones, and being
CC prominently expressed in the liver and the intestine. The methods and
CC compositions of the present invention are useful for identifying a
CC variety of therapeutically useful compounds used in the treatment of a
CC wide variety of indications, such as Cushing's syndrome which leads to
CC obesity, fatigue, hypertension, oedema and osteoporosis, virilism and
CC hirsutism in females due to overproduction of testosterone, androgen
CC excess due to polycystic ovarian syndrome, enzymatic defects which lead
CC to accumulation of specific steroids, and ameliorate the effect of
CC substances in the diet and/or environment which act as endocrine
CC disruptors. This sequence represents steroid hydroxylase rCY3A1 putative
CC SXR (steroid X receptor) response element
XX
SQ Sequence 25 BP; 8 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAGACAGTTCATGAAGTTCATCTAC 25
||||||||||||||||||||||||||
Db 1 TAGACAGTTCATGAAGTTCATCTAC 25

RESULT 5
ACD40529
ID ACD40529 standard; DNA; 25 BP.
XX
AC ACD40529;
XX
DT 04-SEP-2003 (first entry)
XX
DE Rat steroid hydroxylase rCYP3A1 SXR response element.
XX
KW SXR; steroid and xenobiotic receptor; ds; retinoid X receptor; steroid;
KW steroid inducible P450; xenobiotic; homeostasis; drug interaction;
KW Cushing's syndrome; virilism; hirsutism; polycystic ovarian syndrome;
KW prostate cancer; 21-hydroxylase deficiency; 17-hydroxylase deficiency;
KW 3beta-hydroxysteroid dehydrogenase deficiency; colorectal cancer;
KW breast cancer; 11beta-hydroxylase deficiency; steroid toxicity; rat;
KW response element.
XX
OS Rattus sp.
XX
PN US2003044888-A1.
XX
PD 06-MAR-2003.
XX
PF 08-JAN-1999; 99US-00227718.
XX
PR 09-JAN-1998; 98US-00005286.
XX
PA (EVAN/) EVANS R M.
PA (BLUM/) BLUMBERG B.
XX
PI Evans RM, Blumberg B;
XX
DR WPI; 2003-503491/47.
XX
PT New steroid and xenobiotic receptor polypeptides, useful in mediating the
PT physiological effects of steroids and xenobiotics, particularly when
PT combinations of the compounds disrupt homeostatis or cause drug
PT interaction.
XX
PS Disclosure; Fig 6A; 41pp; English.
XX
CC The invention relates to a receptor polypeptide or its functional
CC fragment. The polypeptide forms a heterodimer with retinoid X receptor,
CC binds to a direct or inverted repeat response element motif based on the
CC half site AGTTCA, activates transcription through response elements found
CC in steroid inducible P450 genes in response to a wide variety of natural
CC and synthetic steroid hormones and is prominently expressed in the liver
CC and the intestine. The receptor polypeptides are useful in mediating the
CC physiological effects of steroids and xenobiotics, particularly when
CC combinations of the compounds disrupt homeostatis or cause drug
CC interaction. The receptor polypeptides are also useful in monitoring
CC total steroid levels and inducing the expression of genes encoding
CC xenobiotic metabolising enzymes. The steroid and/or xenobiotic compounds

```
CC   are useful for treating a disease state e.g. Cushing's syndrome, virilism
CC   and hirsutism in females, polycystic ovarian syndrome, 21-hydroxylase
CC   deficiency, 11beta-hydroxylase deficiency, 3beta- hydroxysteroid
CC   dehydrogenase deficiency, 17-hydroxylase deficiency, or breast,
CC   colorectal or prostate cancers. The methods are useful for preventing
CC   steroid toxicity in a subject undergoing treatment of a disease state, or
CC   for slowing clearance of a therapeutic steroid or xenobiotic from a
CC   subject. The present sequence represents a rat steroid and xenobiotic
CC   receptor SXR response element
XX
SQ   Sequence 25 BP; 8 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 25;  DB 9;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 0.14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db          1 TAGACAGTTCATGAAGTTCATCTAC 25


RESULT 6
AAD50114
ID   AAD50114 standard; DNA; 25 BP.
XX
AC   AAD50114;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Rat CYP3A1 steriod hydroxylase SXR response element, DR-3.
XX
KW   Steroid and xenobiotic receptor; SXR; expression system; homeostasis;
KW   steriod hydroxylase; rat; direct repeat; DR; ds.
XX
OS   Rattus sp.
XX
PN   WO200286063-A2.
XX
PD   31-OCT-2002.
XX
PF   16-APR-2002; 2002WO-US012161.
XX
PR   20-APR-2001; 2001US-00840008.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Evans RM;
XX
DR   WPI; 2003-093112/08.
XX
PT   Xenobiotic compound modulated expression systems, useful for modulating
PT   metabolism of one or more endogenous steroids or xenobiotics to establish
PT   homeostasis.
XX
PS   Disclosure; Page 37; 85pp; English.
XX
CC   The invention relates to an expression system which comprises at least
CC   one steroid and xenobiotic receptor (SXR) response element operably
CC   linked to at least one gene and a nuclear receptor which responds to
CC   xenobiotic compounds. Methods of the invention are useful for producing
CC   intracellular receptor target protein in a cell. The methods are also
CC   useful for modulating the physiological response to elevated levels of
CC   steroid and/or xenobiotic compounds to establish homeostasis. The present
CC   sequence is rat CYP3A1 steriod hydroxylase SXR response element, direct
CC   repeat (DR)
XX
SQ   Sequence 25 BP; 8 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 25;  DB 10;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 0.14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db          1 TAGACAGTTCATGAAGTTCATCTAC 25


RESULT 7
ADW22213/c
ID   ADW22213 standard; cDNA; 2073 BP.
XX
AC   ADW22213;
XX
DT   10-MAR-2005  (first entry)
XX
DE   Rat hepatotoxicity marker gene, SEQ:592.
XX
KW   Toxicology screening; drug screening; gene expression;
KW   expression profile; hepatotoxicity; drug-induced; hepatitis;
KW   liver disease; gastrointestinal disease; gene; ss.
XX
OS   Rattus norvegicus.
XX
PN   EP1344834-A2.
XX
PD   17-SEP-2003.
XX
PF   04-MAR-2003; 2003EP-00004810.
XX
PR   14-MAR-2002; 2002EP-00005336.
PR   17-JUL-2002; 2002EP-00015657.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI   Boess F,  Suter-Dick L,  Wolf D;
XX
DR   WPI; 2003-723475/69.
DR   EMBL; X62086.
XX
PT   Predicting toxicity of compounds, useful in development of safe drugs, by
PT   measuring the differential expression of specific genes in cells exposed
PT   to test compounds.
XX
PS   Claim 2; SEQ ID NO 592; 895pp; English.
XX
CC   The invention relates to methods of predicting at least one toxic effect
CC   (or toxicity progression or the mechanism of toxicity) of a compound. The
CC   methods involve detecting the level of expression of at least one of a
CC   set of 680 genes ADW21622-ADW22301 or at least one of a set of 17 genes
CC   (including ADW22362, ADW22414 and ADW22481-ADW22483) in a tissue or cell
CC   exposed to the compound, and determining whether the gene is
CC   differentially expressed compared with a control tissue or cell.
CC   Differential expression of the gene in the presence of the compound is
CC   indicative of a toxic effect, of toxicity progression or of a specific
CC   mechanism of toxicity. The toxic effect is especially hepatotoxicity,
CC   particularly hepatitis, liver necrosis, protein adduct formation or fatty
CC   liver. The invention also relates to sets of primers and probes specific
CC   for at least two genes selected from ADW21622-ADW22301; solid supports
CC   (e.g., DNA chips) and kits containing the probes; and a database
CC   containing DNA sequence information and expression information for at
CC   least two of the 680 genes from hepatotoxin-exposed tissues. The
CC   invention is based on the determination of global changes in gene
CC   expression in tissues or cells exposed to known toxins, particularly
CC   hepatotoxins, and the identification of individual genes (toxicity
CC   markers) that are differentially expressed on toxin exposure. The changes
CC   in gene expression can be characteristic of different mechanisms of
CC   hepatotoxicity mediated by various classes of compounds. Such compounds
CC   include: direct acting compounds which cause damage to macromolecules,
CC   especially proteins and lipids by directly interacting with them;
CC   steatotic compounds which cause an accumulation of fat in the liver; and
CC   cholestatic compounds which impair bile flow or bile acid transport,
CC   resulting in jaundice. The methods of the invention are useful in
CC   toxicology screening for predicting the toxic effects (especially
CC   hepatotoxic effects) of compounds for the development of safer drugs.
CC   Sequences ADW21622-ADW22301 represent specifically claimed hepatotoxicity
CC   marker genes of rat origin whose expression is altered on exposure to
```

```
CC   hepatotoxins.
XX
SQ   Sequence 2073 BP; 471 A; 500 C; 474 G; 628 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 25;  DB 11;  Length 2073;
  Best Local Similarity  100.0%;  Pred. No. 0.32;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||||||||||||
Db       1284 TAGACAGTTCATGAAGTTCATCTAC 1260


RESULT 8
AAZ07991
ID   AAZ07991 standard; DNA; 31 BP.
XX
AC   AAZ07991;
XX
DT   17-JAN-2000  (first entry)
XX
DE   Oligo containing CYP3A1 DR3 PXRE.
XX
KW   Human; nuclear receptor; pregnane X receptor; PXR; CYP; CYP3A4;
KW   cytochrome P-450 mono-oxygenase; drug interaction; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9948915-A1.
XX
PD   30-SEP-1999.
XX
PF   26-MAR-1999;   99WO-US006737.
XX
PR   27-MAR-1998;   98US-0079593P.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kliewer SA,  Willson TM;
XX
DR   WPI; 1999-601202/51.
XX
PT   New human pregnane X receptor, used to identify specific modulators and
PT   agents that induce expression of cytochrome P-450 mono-oxygenase.
XX
PS   Example; Page 20; 69pp; English.
XX
CC   The invention provides an isolated human nuclear receptor (designated
CC   pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase
CC   (CYP) promoter. The hPXR is used to identify: its specific modulators,
CC   and compounds that induce CYP3A4 expression (i.e. to identify drug
CC   interactions, since CYP3A4 is involved in many biotransformations of
CC   drugs). The modulators are potentially useful for: associating particular
CC   diseases and conditions with PXR and for treating such conditions.
CC   Antibodies raised against hPXR can be used for determination and
CC   purification of hPXR. The present sequence represents a double stranded
CC   oligo containing CYP3A1 DR3 PXRE
XX
SQ   Sequence 31 BP; 10 A; 6 C; 6 G; 9 T; 0 U; 0 Other;

  Query Match            92.0%;  Score 23;  DB 2;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 1.2;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTA 24
              |||||||||||||||||||||||
Db          5 AGACAGTTCATGAAGTTCATCTA 27


RESULT 9
ABA91215
ID   ABA91215 standard; DNA; 31 BP.
XX
AC   ABA91215;
XX
DT   04-APR-2002  (first entry)
XX
DE   CYP3A1 DR3 pregnane X receptor response element.
XX
KW   Pregnane X; receptor; hPXR; rat; cytochrome P450 monooxygenase; CYP3A1;
KW   liver; PXRE; ss.
XX
OS   Rattus sp.
XX
PN   WO200197856-A2.
XX
PD   27-DEC-2001.
XX
PF   21-JUN-2001; 2001WO-IB001629.
XX
PR   21-JUN-2000; 2000US-00598267.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kliewer SA,  Jones SA,  Willson TM;
XX
DR   WPI; 2002-139767/18.
XX
PT   Compound that induces cytochrome P-450 monooxygenase 3A4 gene expression
PT   for treating cholestatic liver disease comprising administering compound
PT   identified by determining binding of test compound to human pregnane X
PT   receptor.
XX
PS   Example 1; Page 22; 63pp; English.
XX
CC   The present sequence is that of the rat cytochrome P450 monooxygenase 3A1
CC   gene (CYP3A1) DR3 pregnane X receptor response element (PXRE). 4 Copies
CC   of this sequence were inserted into the BamHI site of pBLCAT2 to create
CC   reporter plasmid (DR3)4-tk-CAT. This was used in CV1 transfection assays
CC   to demonstrate that novel human pregnane X receptor (hPXR, see AAM50624)
CC   is a functional nuclear receptor that is activated by dexamethasone-t-
CC   butylacetate, a known mPXR1 ligand. The oligonucleotide was also used in
CC   band shift assays, which showed that hPXR binds efficiently to the CYP3A4
CC   IR6 PXRE as a heterodimer with the 6-cis retinoic acid receptor, and that
CC   hPXR and mPXR1 have very similar DNA binding profiles. The invention
CC   provides nucleic acids encoding hPXR, expression vectors, host cells, and
CC   methods of using the receptor-encoding sequences to screen for compounds
CC   capable of modulating CYP (e.g. CYP3A4) gene expression. Such compounds
CC   are useful for treating cholestatic liver disease (claimed), such as
CC   primary biliary cirrhosis, primary sclerosing cholangitis, autoimmune
CC   hepatitis with cholestatic features, autoimmune cholangitis, cholestasis
CC   of pregnancy, paediatric cholestatic syndromes, and drug-induced
CC   cholestasis
XX
SQ   Sequence 31 BP; 10 A; 6 C; 6 G; 9 T; 0 U; 0 Other;

  Query Match            92.0%;  Score 23;  DB 6;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 1.2;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTA 24
              |||||||||||||||||||||||
Db          5 AGACAGTTCATGAAGTTCATCTA 27


RESULT 10
AAZ07996
ID   AAZ07996 standard; DNA; 32 BP.
XX
AC   AAZ07996;
XX
DT   17-JAN-2000  (first entry)
XX
DE   Radiolabeled probe CYP3A1 DR3.
XX
```

KW Human; nuclear receptor; pregnane X receptor; PXR; CYP; CYP3A4;
KW cytochrome P-450 mono-oxygenase; drug interaction; probe; ss.
XX
OS Synthetic.
XX
PN WO9948915-A1.
XX
PD 30-SEP-1999.
XX
PF 26-MAR-1999; 99WO-US006737.
XX
PR 27-MAR-1998; 98US-0079593P.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kliewer SA, Willson TM;
XX
DR WPI; 1999-601202/51.
XX
PT New human pregnane X receptor, used to identify specific modulators and
PT agents that induce expression of cytochrome P-450 mono-oxygenase.
XX
PS Example; Page 22; 69pp; English.
XX
CC The invention provides an isolated human nuclear receptor (designated
CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase
CC (CYP) promoter. The hPXR is used to identify: its specific modulators,
CC and compounds that induce CYP3A4 expression (i.e. to identify drug
CC interactions, since CYP3A4 is involved in many biotransformations of
CC drugs). The modulators are potentially useful for: associating particular
CC diseases and conditions with PXR and for treating such conditions.
CC Antibodies raised against hPXR can be used for determination and
CC purification of hPXR. Sequences AAZ07993-996 represent radiolabeled
CC probes or competitors used in band shift assays
XX
SQ Sequence 32 BP; 10 A; 6 C; 7 G; 9 T; 0 U; 0 Other;

Query Match 92.0%; Score 23; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGACAGTTCATGAAGTTCATCTA 24
||||||||||||||||||||||||
Db 6 AGACAGTTCATGAAGTTCATCTA 28

RESULT 11
AAZ40699/c
ID AAZ40699 standard; DNA; 21 BP.
XX
AC AAZ40699;
XX
DT 14-MAR-2000 (first entry)
XX
DE Rat CYP3A23 enhancer nuclear receptor response element CYP3A23 DR3.
XX
KW Transcriptional enhancer; cytochrome P450; CYP3A4; genetic analysis;
KW drug metabolism; prostatic cancer; xenobiotic; therapeutic drug; ss;
KW genetic switch; transgene activation; nuclear receptor response element.
XX
OS Rattus sp.
XX
PN WO9961622-A1.
XX
PD 02-DEC-1999.
XX
PF 20-MAY-1999; 99WO-AU000381.
XX
PR 21-MAY-1998; 98AU-00003628.
XX
PA (UNSY) UNIV SYDNEY.
XX
PI Liddle C, Goodwin BJ;
XX
DR WPI; 2000-072626/06.
XX
PT New nucleic acid containing a transcriptional enhancer of cytochrome P450
PT CYP3A4, used to identify xenobiotics that induce cytochrome expression.
XX
PS Disclosure; Page 11; 38pp; English.
XX
CC The invention relates to an isolated nucleic acid (I) containing a
CC transcriptional enhancer of the production or expression of cytochrome
CC P450 CYP3A4. (I), or its fragments, are useful in genetic analysis,
CC particularly for: detecting allelic variants of (I), for predicting drug
CC metabolism and susceptibility to disease (particularly prostatic cancer),
CC and for analysis of the effect of allelic variations on CYP3A4
CC transcription and expression. Assay systems that include (I) linked to a
CC reporter sequence are used to screen xenobiotics (therapeutic drugs) for
CC ability to induce CYP3A4 expression in cells or animals. Candidate drugs
CC that induce CYP3A4 will: (a) have reduced in vivo lifetime, since they
CC will be metabolized by CYP3A4, or (b) may increase metabolism and/or
CC elimination of co-administered drugs. Such compounds should be discarded
CC in favor of non-inducing candidates. Also, induction of CYP3A4 can be
CC used to accelerate metabolism of xenobiotic toxins or endogenous CYP3A4
CC substrates, while inhibition of CYP3A4 can be used to overcome
CC interactions with drugs. (I) is also potentially useful as a genetic
CC switch, e.g. for activating a transgene in the liver. Identification of
CC structural motifs associated with CYP3A4 induction may be used for
CC rational drug design. The present sequence represents a nuclear receptor
CC response element in the proximal 5' flanking region of the rat CYP3A23
XX
SQ Sequence 21 BP; 7 A; 4 C; 4 G; 6 T; 0 U; 0 Other;

Query Match 84.0%; Score 21; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GACAGTTCATGAAGTTCATCT 23
|||||||||||||||||||||
Db 21 GACAGTTCATGAAGTTCATCT 1

RESULT 12
AAX89082
ID AAX89082 standard; DNA; 25 BP.
XX
AC AAX89082;
XX
DT 14-SEP-1999 (first entry)
XX
DE Putative SXR response element DR-3 containing fragment rCYP3A2.
XX
KW Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;
KW retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;
KW phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;
KW breast cancer; osteoporosis; Cushing syndrome; virilism; hirsutism;
KW polycystic ovarian disease; cancer; colorectal; prostatic; ss.
XX
OS Homo sapiens.
XX
PN WO9935246-A1.
XX
PD 15-JUL-1999.
XX
PF 08-JAN-1999; 99WO-US000490.
XX
PR 09-JAN-1998; 98US-00005286.
XX
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Evans RM, Blumberg B;
XX
DR WPI; 1999-419349/35.
XX
PT New steroid and xenobiotic receptor, used to identify modulators for

```
PT   controlling metabolism of steroids and xenobiotics, e.g. reducing their
PT   toxicity.
XX
PS   Disclosure; Fig 6A; 83pp; English.
XX
CC   The invention relates to a novel nuclear receptor polypeptide, designated
CC   SXR (steroid and xenobiotic receptor). SXR (i) forms a heterodimer with
CC   retinoid X receptor (RXR), (ii) binds to a direct or inverted repeat
CC   response element motif based on the half-site AGTTCA, (iii) activates
CC   transcription through response elements present in steroid-inducible P450
CC   genes, in response to a wide variety of natural and synthetic steroid
CC   hormones and (iv) is prominently expressed in liver and intestine. SXR
CC   regulates expression of catabolic enzymes, in response to many different
CC   steroids, and thus affects metabolism. SXR is a broad specificity, low-
CC   affinity receptor for reducing excessive levels of steroids in the
CC   circulation (see AAX89090 for detailed uses of SXR polypeptide).
CC   Sequences AAX89081-89 represent fragments from various steroid and
CC   xenobiotic inducible P450 enzymes containing putative SXR response
CC   elements
XX
SQ   Sequence 25 BP; 9 A; 5 C; 3 G; 8 T; 0 U; 0 Other;

  Query Match             80.8%;  Score 20.2;  DB 2;  Length 25;
  Best Local Similarity   88.0%;  Pred. No. 22;
  Matches   22;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              ||  |||||||| ||||||||||||
Db          1 TAAGCAGTTCATAAAGTTCATCTAC 25


RESULT 13
AAH25491
ID   AAH25491 standard; DNA; 25 BP.
XX
AC   AAH25491;
XX
DT   22-AUG-2001  (first entry)
XX
DE   Steroid-activated nuclear receptor putative response element.
XX
KW   Steroid-activated nuclear receptor; steroid and xenobiotic receptor; SXR;
KW   retinoid X receptor; RXR; transcription; response element;
KW   steroid inducible P450 gene; steroid hormone; Cushing's syndrome;
KW   virulism; hirsutism; polycystic ovarian syndrome; hypertension; ss.
XX
OS   Unidentified.
XX
PN   WO200142290-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033473.
XX
PR   09-DEC-1999;   99US-00458366.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Evans RM,  Blumberg B,  Xie W;
XX
DR   WPI; 2001-381637/40.
XX
PT   Novel steroid-activated nuclear receptor useful as sensor for xenobiotic
PT   compounds and/or steroids and whose modulators are useful for modulating
PT   metabolism of steroid or xenobiotic compounds.
XX
PS   Disclosure; Page 23; 64pp; English.
XX
CC   The present sequence represents a putative response element for a steroid
CC   -activated nuclear receptor, termed steroid and xenobiotic receptor
CC   (SXR). The response element is identified in steroid hydroxylase CYP3A2.
CC   The SXR polypeptide is capable of forming a heterodimer with retinoid X
CC   receptor (RXR), activating transcription through response elements found
CC   in steroid inducible P450 genes in response to a variety of natural and
CC   synthetic steroid hormones and prominently expressed in liver and
CC   intestine. SXR binds to a direct or inverted repeat response element
CC   motif based on the half site AGTTCA. SXR is useful for identifying
CC   compounds which are agonists or which activate the receptor. The
CC   compounds identified are useful for treating a wide variety of conditions
CC   such as Cushing's syndrome, virulism and hirsutism, androgen excess due
CC   to polycystic ovarian syndrome and enzymatic defects which leads to
CC   accumulation of steroids, resulting in hypertension and aberrant
CC   development of secondary sexual characteristics in both sexes. Transgenic
CC   animals which express human SXR serve as models for human response to
CC   various agents which potentially impact P450-dependent metabolic
CC   processes
XX
SQ   Sequence 25 BP; 9 A; 5 C; 3 G; 8 T; 0 U; 0 Other;

  Query Match             80.8%;  Score 20.2;  DB 5;  Length 25;
  Best Local Similarity   88.0%;  Pred. No. 22;
  Matches   22;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              ||  |||||||| ||||||||||||
Db          1 TAAGCAGTTCATAAAGTTCATCTAC 25


RESULT 14
ABZ58305
ID   ABZ58305 standard; DNA; 25 BP.
XX
AC   ABZ58305;
XX
DT   28-APR-2003  (first entry)
XX
DE   Direct repeat-3 in rat cytochrome P450-3A2 gene.
XX
KW   Steroid xenobiotic receptor; SXR; receptor; cytochrome-P450; rat;
KW   steroid; xenobiotic; antidote; detoxification; ds.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   repeat_region    6. .20
FT                    /*tag=  a
FT                    /rpt_type= DIRECT
FT                    /function= "Response element"
FT   repeat_unit      6. .11
FT                    /*tag=  b
FT   repeat_unit      15. .20
FT                    /*tag=  c
XX
PN   WO2003005812-A2.
XX
PD   23-JAN-2003.
XX
PF   09-JUL-2002; 2002WO-US021800.
XX
PR   09-JUL-2001; 2001US-0304388P.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Evans R,  Xie W;
XX
DR   WPI; 2003-221630/21.
XX
PT   Modulating the metabolism of steroids and xenobiotics with a UGT
PT   modulator, useful for modifying the physiological response to and/or
PT   efficient detoxification of harmful steroids and/or xenobiotic compounds.
XX
PS   Disclosure; Page 26; 51pp; English.
XX
CC   The present sequence is a direct repeat-3 (DR-3) response element for the
CC   steroid xenobiotic receptor (SXR/PXR) identified in the rat cytochrome
CC   P450-3A2 gene. A database search showed that putative SXR response
```

CC elements are found in genes encoding steroid hydroxylases, P450
CC oxidoreductase, and glucuronosyl transferase. SXR is a broad specificity,
CC low affinity, steroid-activated receptor. The present invention relates
CC to modulation of metabolism of steroids and xenobiotics. Nuclear
CC receptors including SXR and constitutively active receptor (CAR) are
CC characterised as xenosensors regulating expression of P450 genes. The
CC ability of this group of receptors to regulate expression of UDP-
CC glucuronosyl transferase (UGT) in response to steroids and/or xenobiotics
CC provides novel approaches for direct regulation/activation of a
CC glucuronidation pathway, thereby providing methods to achieve physiologic
CC homeostasis with respect to steroids and/or xenobiotics. SXR and CAR
CC regulation of UGT represents the first evidence of receptors that can
CC transduce/transactivate both Phase I and Phase II adaptive hepatic
CC response. A claimed method for modulating the metabolism or clearance of
CC steroid and/or xenobiotic compounds involves administering a modulator of
CC UGT. The modulator can be a nucleic acid, protein and/or chemical
CC compound which binds a UGT direct repeat or inverted repeat response
CC element, or which activates UGT glucuronidation of steroids or
CC xenobiotics. A claimed method for inducing expression of steroid
CC hydroxylase comprises activating SXR/PXR and/or CAR
XX
SQ Sequence 25 BP; 9 A; 5 C; 3 G; 8 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 8; Length 25;
Best Local Similarity 88.0%; Pred. No. 22;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TAGACAGTTCATGAAGTTCATCTAC 25
|| |||||||| ||||||||||||
Db 1 TAAGCAGTTCATAAAGTTCATCTAC 25

RESULT 15
ACD27770
ID ACD27770 standard; DNA; 25 BP.
XX
AC ACD27770;
XX
DT 18-SEP-2003 (first entry)
XX
DE Steroid hydroxylase rCYP3A2, putative SXR response element.
XX
KW Human; steroid X receptor; SXR; retinoid X receptor; RXR;
KW steroid inducible P450 gene; Cushing's syndrome; obesity; fatigue;
KW hypertension; oedema; osteoporosis; virilism; hirsutism; androgen excess;
KW polycystic ovarian syndrome; steroids accumulation; endocrine disruptor;
KW steroid hydroxylase; rCYP3A2; SXR response element; ds.
XX
OS Unidentified.
XX
PN US2003064430-A1.
XX
PD 03-APR-2003.
XX
PF 09-JAN-1998; 98US-00005286.
XX
PR 09-JAN-1998; 98US-00005286.
XX
PA (EVAN/) EVANS R M.
PA (BLUM/) BLUMBERG B.
XX
PI Evans RM, Blumberg B;
XX
DR WPI; 2003-540786/51.
XX
PT New steroid-activated nuclear receptor polypeptide that heterodimerizes
PT with retinoid X receptor, useful for identifying therapeutic compounds
PT for the treatment of Cushing's syndrome, virilism and hirsutism, and
PT androgen excess.
XX
PS Disclosure; Page 4; 23pp; English.
XX
CC The invention describes a new receptor polypeptide (I) or its functional
CC fragments. The method comprises forming a heterodimer with retinoid X
CC receptor (RXR), binding to a direct or inverted repeat response element
CC motif based on the half site AGTTCA, activating transcription through
CC response elements found in steroid inducible P450 genes in response to a
CC wide variety of natural and synthetic steroid hormones, and being
CC prominently expressed in the liver and the intestine. The methods and
CC compositions of the present invention are useful for identifying a
CC variety of therapeutically useful compounds used in the treatment of a
CC wide variety of indications, such as Cushing's syndrome which leads to
CC obesity, fatigue, hypertension, oedema and osteoporosis, virilism and
CC hirsutism in females due to overproduction of testosterone, androgen
CC excess due to polycystic ovarian syndrome, enzymatic defects which lead
CC to accumulation of specific steroids, and ameliorate the effect of
CC substances in the diet and/or environment which act as endocrine
CC disruptors. This sequence represents steroid hydroxylase rCY3A2 putative
CC SXR (steroid X receptor) response element
XX
SQ Sequence 25 BP; 9 A; 5 C; 3 G; 8 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 9; Length 25;
Best Local Similarity 88.0%; Pred. No. 22;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TAGACAGTTCATGAAGTTCATCTAC 25
|| |||||||| ||||||||||||
Db 1 TAAGCAGTTCATAAAGTTCATCTAC 25

Search completed: February 6, 2006, 14:42:13
Job time : 290 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 14:42:31 ; Search time 2323 Seconds
(without alignments)
503.520 Million cell updates/sec

Title: US-10-081-555C-3
Perfect score: 25
Sequence: 1 tagacagttcatgaagttcatctac 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20.2	80.8	715	5	BQ991522	BQ991522 QGF23B15.
c	2	19.8	79.2	756	2	BE914342	BE914342 601669311
c	3	19.8	79.2	977	9	CC215499	CC215499 CH261-189
c	4	19.4	77.6	564	5	BU216276	BU216276 603755845
c	5	19.4	77.6	571	1	AL587998	AL587998
c	6	19.4	77.6	640	5	BU425734	BU425734 603231044
c	7	19.4	77.6	700	7	CN227903	CN227903 RJB097G11
c	8	19.4	77.6	722	5	BU298821	BU298821 603610915
c	9	19.4	77.6	918	5	BU459792	BU459792 603367405
	10	19.2	76.8	123	1	AA640771	AA640771 nu02f01.s
	11	19.2	76.8	556	11	DE078766	DE078766 Oryzias l
	12	19.2	76.8	564	9	AZ023470	AZ023470 RPCI-23-2
	13	19.2	76.8	680	8	CX302184	CX302184 C08018A08
c	14	19.2	76.8	770	7	CK000827	CK000827 AGENCOURT
	15	19.2	76.8	892	8	DN023192	DN023192 JGI_CAAR3
	16	19.2	76.8	925	8	DN017667	DN017667 JGI_CAAR6
c	17	18.8	75.2	465	1	AJ696924	AJ696924
c	18	18.8	75.2	473	6	CB222230	CB222230 1IL25E10
c	19	18.8	75.2	547	1	AV663133	AV663133 AV663133
c	20	18.8	75.2	555	3	BM105335	BM105335 508767 MA
c	21	18.8	75.2	572	2	BE665178	BE665178 153287 MA
c	22	18.8	75.2	572	3	BI679920	BI679920 457214 MA
c	23	18.8	75.2	576	6	CD288024	CD288024 2_L7.abd
c	24	18.8	75.2	583	7	CN434008	CN434008 BE030006A
c	25	18.8	75.2	594	6	CB452671	CB452671 707552 MA
	26	18.8	75.2	629	1	AV647972	AV647972 AV647972
c	27	18.8	75.2	657	7	CN792153	CN792153 4126982 B
	28	18.8	75.2	663	3	BJ001765	BJ001765 BJ001765
c	29	18.8	75.2	673	1	AV609393	AV609393 AV609393
c	30	18.8	75.2	680	7	CK947925	CK947925 4072661 B
c	31	18.8	75.2	685	8	DN533820	DN533820 1366725 M
c	32	18.8	75.2	706	4	AY069009	AY069009 Schmidtea
c	33	18.8	75.2	735	10	CG049226	CG049226 PUIFH41TB
	34	18.8	75.2	747	3	BJ677231	BJ677231 BJ677231
c	35	18.8	75.2	756	7	CO888306	CO888306 BovGen_16
c	36	18.8	75.2	780	8	DN823975	DN823975 LB0011.C2
	37	18.8	75.2	830	10	CZ539066	CZ539066 SRAA-aad2
c	38	18.8	75.2	833	1	AM005097	AM005097 AM005097
c	39	18.8	75.2	839	2	BI144624	BI144624 602909911
	40	18.8	75.2	869	11	CR808491	CR808491 GR0AAA31B
c	41	18.8	75.2	870	1	AM011102	AM011102 AM011102
c	42	18.8	75.2	933	9	CC399823	CC399823 PUHGS37TD
	43	18.6	74.4	206	7	CO651738	CO651738 cct07_H01
c	44	18.6	74.4	420	9	AQ300850	AQ300850 HS_2217_B
	45	18.6	74.4	427	1	AA627035	AA627035 MBAFCW2G0

ALIGNMENTS

RESULT 1
BQ991522
LOCUS BQ991522 715 bp mRNA linear EST 21-AUG-2002
DEFINITION QGF23B15.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone QGF23B15, mRNA sequence.
ACCESSION BQ991522
VERSION BQ991522.1 GI:22411057
KEYWORDS EST.
SOURCE Lactuca sativa
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.
REFERENCE 1 (bases 1 to 715)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
singleton, see http://cgpdb.ucdavis.edu/ for details.
Plate: QGF23 row: B column: 15.
FEATURES Location/Qualifiers
source 1. .715
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGF23B15"
/lab_host="E.coli"
/clone_lib="QG_EFGHJ lettuce serriola"
/note="Vector: pBRcDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated,

directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=flowers post-fertilized
TAG_LIB=QG_EFGHJ lettuce serriola
TAG_SEQ=TGCCATCGGG"

ORIGIN

Query Match 80.8%; Score 20.2; DB 5; Length 715;
Best Local Similarity 88.0%; Pred. No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||| |||||||||||||||| | |
Db        232 TAGAGAGTTCATGAAGTTCATTTTC 256
```

RESULT 2
BE914342/c
LOCUS BE914342 756 bp mRNA linear EST 29-SEP-2000
DEFINITION 601669311F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969026 5', mRNA sequence.
ACCESSION BE914342
VERSION BE914342.1 GI:10412869
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 756)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM9146 row: f column: 03
High quality sequence stop: 713.

FEATURES Location/Qualifiers
source 1. .756
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969026"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 79.2%; Score 19.8; DB 2; Length 756;
Best Local Similarity 91.3%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 GACAGTTCATGAAGTTCATCTAC 25
              ||| ||||||||||||||| |||
Db        509 GACGGTTCATGAAGTTCATGTAC 487
```

RESULT 3
CC215499/c
LOCUS CC215499 977 bp DNA linear GSS 12-MAY-2003
DEFINITION CH261-189L2_Sp6.1 CH261 Gallus gallus genomic clone CH261-189L2, genomic survey sequence.
ACCESSION CC215499
VERSION CC215499.1 GI:30534167
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 977)
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 52
High quality sequence stop: 771.

FEATURES Location/Qualifiers
source 1. .977
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-189L2"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI; CH261 Female Chicken library - for library and clone ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 79.2%; Score 19.8; DB 9; Length 977;
Best Local Similarity 91.3%; Pred. No. 6e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 GACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||| |||| ||||
Db        709 GACAGTTCATGAATTTCACCTAC 687
```

RESULT 4
BU216276/c
LOCUS BU216276 564 bp mRNA linear EST 25-NOV-2002
DEFINITION 603755845F1 CSEQCHN04 Gallus gallus cDNA clone ChEST666j12 5', mRNA sequence.
ACCESSION BU216276
VERSION BU216276.1 GI:25395849
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 564)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)

```
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES              Location/Qualifiers
     source           1. .564
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="White Leghorn, Hisex"
                      /db_xref="taxon:9031"
                      /clone="ChEST666j12"
                      /tissue_type="whole embryo"
                      /dev_stage="20-21"
                      /lab_host="DH10B"
                      /clone_lib="CSEQCHN04"
                      /note="Organ: whole embryo; Vector: pBluescript II KS(+);
                      Site_1: EcoRI; Site_2: Not1; This normalized library was
                      constructed from 1 million independent clones. cDNA
                      synthesis was initiated using an oligo(dT) primer, using
                      methylated C in the first strand synthesis reaction.
                      Following this first strand reaction, double-stranded cDNA
                      was blunted, ligated to NotI adapters, digested with
                      EcoRI, size-selected, and cloned into the NotI and EcoRI
                      compatible sites of a custom modified MCS of the
                      pBluescript (KS+) vector. The library was normalized in 2
                      rounds using conditions adapted from Soares et al., PNAS
                      (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                      (1996): 791, except that a significantly longer
                      reannealing hybridization was used."
ORIGIN

  Query Match             77.6%;  Score 19.4;  DB 5;  Length 564;
  Best Local Similarity   95.2%;  Pred. No. 8.3e+02;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATC 22
              | ||||||||||||||||||||
Db        152 ATACAGTTCATGAAGTTCATC 132


RESULT 5
AL587998/c
LOCUS       AL587998                 571 bp    mRNA    linear   EST 02-MAR-2001
DEFINITION  AL587998 BP Chicken Brain Library Gallus gallus cDNA clone
            ROS066C05, mRNA sequence.
ACCESSION   AL587998
VERSION     AL587998.1  GI:13193032
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Murray,F.
  TITLE     BP Chicken Brain Library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Frazer Murray
            Dept. Genomics and Bioinformatics
            Roslin Institute
            Roslin, Midlothian, EH25 9PS, UK
            Tel: +44 (0)131 527 4200
            Fax: +44 (0)131 440 0434
            Email: frazer.murray@bbsrc.ac.uk
            GCGGCCGCTTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetech
            (*6854-
            Seq primer: T7.
FEATURES              Location/Qualifiers
     source           1. .571
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9031"
                      /clone="ROS066C05"
                      /tissue_type="Brain"
                      /dev_stage="Unknown"
                      /lab_host="DH10B"
                      /clone_lib="BP Chicken Brain Library"
                      /note="Vector: pSPORT1; Site_1: Not1; Site_2: Sal1; Cloned
                      unidirectionally. Primer: Oligo dT. 5' adaptor sequence:
                      5' TCGACCTCGAG 3' ; 3' adaptor sequence: 5'
                      GCGGCCGCTTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from
                      Clonetech (*6854-1)"
ORIGIN

  Query Match             77.6%;  Score 19.4;  DB 1;  Length 571;
  Best Local Similarity   95.2%;  Pred. No. 8.3e+02;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATC 22
              | ||||||||||||||||||||
Db        309 ATACAGTTCATGAAGTTCATC 289


RESULT 6
BU425734/c
LOCUS       BU425734                 640 bp    mRNA    linear   EST 29-NOV-2002
DEFINITION  603231044F1 CSEQRBN09 Gallus gallus cDNA clone ChEST226h12 5', mRNA
            sequence.
ACCESSION   BU425734
VERSION     BU425734.1  GI:25918410
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES              Location/Qualifiers
     source           1. .640
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="Layer and broiler"
                      /db_xref="taxon:9031"
                      /clone="ChEST226h12"
                      /sex="Male and female"
                      /tissue_type="Chondrocytes isolated from growth plate
                      cartilage"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="CSEQRBN09"
                      /note="Vector: pBluescript II KS(+); Site_1: EcoRI;
                      Site_2: Not1; This normalized library was constructed from
                      1 million independent clones. cDNA synthesis was initiated
                      using an oligo(dT) primer, using methylated C in the first
                      strand synthesis reaction. Following this first strand
                      reaction, double-stranded cDNA was blunted, ligated to
                      NotI adapters, digested with EcoRI, size-selected, and
                      cloned into the NotI and EcoRI compatible sites of a
                      custom modified MCS of the pBluescript (KS+) vector. The
                      library was normalized in 2 rounds using conditions
                      adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
                      Bonaldo et al., Genome Research 6 (1996): 791, except that
                      a significantly longer reannealing hybridization was
```

used."
ORIGIN

Query Match 77.6%; Score 19.4; DB 5; Length 640;
Best Local Similarity 95.2%; Pred. No. 8.5e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGACAGTTCATGAAGTTCATC 22
| ||||||||||||||||||||
Db 256 ATACAGTTCATGAAGTTCATC 236

RESULT 7
CN227903/c
LOCUS CN227903 700 bp mRNA linear EST 09-APR-2004
DEFINITION RJB097G11.ab1 RJtestis Gallus gallus cDNA 5', mRNA sequence.
ACCESSION CN227903
VERSION CN227903.1 GI:46331647
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 700)
AUTHORS Savolainen,P., Fitzsimmons,C.J., Arvestad,L., Andersson,L. and Lundeberg,J.
TITLE EST analysis of brain and testis cDNA libraries from White leghorn and Red Jungle Fowl
JOURNAL Unpublished (2004)
COMMENT Contact: Peter Savolainen
Department of Biotechnology
Royal Institute of Technology, KTH
SE-106 91 Stockholm, SWEDEN
Tel: +46 (0)8 5537 8481
Fax: +46 (0)8 5537 8335
Email: Peter.Savolainen@biotech.kth.se
Seq primer: M13 reverse primer.
FEATURES Location/Qualifiers
source 1. .700
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Red junglefowl"
/db_xref="taxon:9031"
/sex="male"
/lab_host="ElectroMAX DH10B (Invitrogen)"
/clone_lib="RJtestis"
/note="Organ: testis; Vector: pSPORT-1; Site_1: Hind III; Site_2: EcoRI; The cDNA libraries were created with the Superscript Plasmid System (Invitrogen)."
ORIGIN

Query Match 77.6%; Score 19.4; DB 7; Length 700;
Best Local Similarity 95.2%; Pred. No. 8.6e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGACAGTTCATGAAGTTCATC 22
| ||||||||||||||||||||
Db 104 ATACAGTTCATGAAGTTCATC 84

RESULT 8
BU298821/c
LOCUS BU298821 722 bp mRNA linear EST 27-NOV-2002
DEFINITION 603610915F1 CSEQCHN56 Gallus gallus cDNA clone ChEST600n4 5', mRNA sequence.
ACCESSION BU298821
VERSION BU298821.1 GI:25748457
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 722)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
FEATURES Location/Qualifiers
source 1. .722
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 15I"
/db_xref="taxon:9031"
/clone="ChEST600n4"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN56"
/note="Organ: small intestine; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."
ORIGIN

Query Match 77.6%; Score 19.4; DB 5; Length 722;
Best Local Similarity 95.2%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGACAGTTCATGAAGTTCATC 22
| ||||||||||||||||||||
Db 244 ATACAGTTCATGAAGTTCATC 224

RESULT 9
BU459792/c
LOCUS BU459792 918 bp mRNA linear EST 29-NOV-2002
DEFINITION 603367405F1 CSEQRBN19 Gallus gallus cDNA clone ChEST268j1 5', mRNA sequence.
ACCESSION BU459792
VERSION BU459792.1 GI:25949103
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 918)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology

```
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES              Location/Qualifiers
     source           1. .918
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="Layer"
                      /db_xref="taxon:9031"
                      /clone="ChEST268j1"
                      /sex="Female"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="CSEQRBN19"
                      /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
                      EcoRI; Site_2: Not1; This normalized library was
                      constructed from 1 million independent clones. cDNA
                      synthesis was initiated using an oligo(dT) primer, using
                      methylated C in the first strand synthesis reaction.
                      Following this first strand reaction, double-stranded cDNA
                      was blunted, ligated to NotI adapters, digested with
                      EcoRI, size-selected, and cloned into the NotI and EcoRI
                      compatible sites of a custom modified MCS of the
                      pBluescript (KS+) vector. The library was normalized in 2
                      rounds using conditions adapted from Soares et al., PNAS
                      (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                      (1996): 791, except that a significantly longer
                      reannealing hybridization was used."
ORIGIN

  Query Match             77.6%;  Score 19.4;  DB 5;  Length 918;
  Best Local Similarity   95.2%;  Pred. No. 9e+02;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATC 22
              | ||||||||||||||||||||
Db        228 ATACAGTTCATGAAGTTCATC 208


RESULT 10
AA640771
LOCUS       AA640771                123 bp    mRNA    linear   EST 27-OCT-1997
DEFINITION  nu02f01.s1 NCI_CGAP_Alv1 Homo sapiens cDNA clone IMAGE:1206841,
            mRNA sequence.
ACCESSION   AA640771
VERSION     AA640771.1  GI:2566021
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 123)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
            Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .123
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1206841"
                      /tissue_type="alveolar rhabdomyosarcoma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Alv1"
                      /note="Vector: pAMP10; mRNA made from alveolar
                      rhabdomyosarcoma, cDNA made by oligo-dT priming.
                      Non-directionally cloned. Size-selected on agarose gel,
                      average insert  size 600 bp. Reference:  Krizman et al.
                      (1996) Cancer Research 56:5380-5383."
ORIGIN

  Query Match             76.8%;  Score 19.2;  DB 1;  Length 123;
  Best Local Similarity   87.5%;  Pred. No. 7.9e+02;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||  || |||||  |||
Db         46 AGACAGTTCATCAAATTCATATAC 69


RESULT 11
DE078766
LOCUS       DE078766                556 bp    DNA     linear   GSS 25-MAY-2005
DEFINITION  Oryzias latipes DNA, clone: ola1-200M05.R, genomic survey sequence.
ACCESSION   DE078766
VERSION     DE078766.1  GI:62597988
KEYWORDS    GSS.
SOURCE      Oryzias latipes (Japanese medaka)
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1
  AUTHORS   Fujiyama,A., Toyoda,A., Kuroki,Y. and Sakaki,Y.
  TITLE     BAC end sequences of Ola1 Oryzias latipes Library
  JOURNAL   Published Only in Database (2005)
REFERENCE   2  (bases 1 to 556)
  AUTHORS   Fujiyama,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2005) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:afujiyam@gsc.riken.jp, URL:http://stt.gsc.riken.jp/,
            Tel:81-3-4212-2558, Fax:81-3-3556-1916)
COMMENT     This work was done in collaboration with Takeda, H. (1), Naruse, K.
            (2)
            and Narita, T. (3)
            (1) Department of Biological Science,
            University of Tokyo
            Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
            Phone: +81-3-5841-4431
            Fax: +81-3-5841-4993
            E-mail: htakeda.s.u-tokyo.ac.jp
            (2) Department of Biological Science,
            University of Tokyo
            Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
            Phone: +81-3-5841-4431
            Fax: +81-3-5841-4993
            E-mail: naruse.s.u-tokyo.ac.jp
            (3) Department of Biological Science,
            University of Tokyo
            Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
            Phone: +81-3-5841-4431
            Fax: +81-3-5841-4993
            E-mail: tanarita.s.u-tokyo.ac.jp
            PRIMERS
                 Sequencing : Forward
            LIBRARY
                 Vector     : pKS145
                 R.Site 1   : SacI
```

```
                L.Site 2    : SacI.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Oryzias latipes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8090"
                     /clone="ola1-200M05.R"
                     /sex="male"
                     /cell_type="whole body"
                     /clone_lib="BAC end sequences of Ola1 Oryzias latipes
                     library"
ORIGIN

  Query Match             76.8%;  Score 19.2;  DB 11;  Length 556;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTAC 25
              ||||||| |||| || ||||||||
Db        291 AGACAGTGCATGTAGGTCATCTAC 314


RESULT 12
AZ023470
LOCUS       AZ023470                564 bp    DNA     linear   GSS 25-FEB-2000
DEFINITION  RPCI-23-276F15.TV RPCI-23 Mus musculus genomic clone
            RPCI-23-276F15, genomic survey sequence.
ACCESSION   AZ023470
VERSION     AZ023470.1  GI:7098854
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-276F15.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 276  row: F  column: 15
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .564
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-276F15"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match             76.8%;  Score 19.2;  DB 9;  Length 564;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTA 24
              |||||||| |||||||  ||||||
Db        411 TAGACAGTACATGAAGAACATCTA 434


RESULT 13
CX302184
LOCUS       CX302184                680 bp    mRNA    linear   EST 06-MAY-2005
DEFINITION  C08018A08SK PhyRootSw1 Citrus sinensis cDNA clone C08018A08, mRNA
            sequence.
ACCESSION   CX302184
VERSION     CX302184.1  GI:63071038
KEYWORDS    EST.
SOURCE      Citrus sinensis
  ORGANISM  Citrus sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Forment,J., Gadea,J., Huerta,L., Abizanda,L., Agusti,J., Alamar,S.,
            Alos,E., Andres,F., Arribas,R., Beltran,J.P., Berbel,A.,
            Blazquez,M.A., Brumos,J., Canas,L.A., Cercos,M.,
            Colmenero-Flores,J.M., Conesa,A., Estables,B., Gandia,M.,
            Garcia-Martinez,J.L., Gimeno,J., Gisbert,A., Gomez,G.,
            Gonzalez-Candelas,L., Granell,A., Guerri,J., Lafuente,M.T.,
            Madueno,F., Marcos,J.F., Marques,M.C., Martinez,F.,
            Martinez-Godoy,M.A., Miralles,S., Moreno,P., Navarro,L., Pallas,V.,
            Perez-Amador,M.A., Perez-Valle,J., Pons,C., Rodrigo,I.,
            Rodriguez,P.L., Royo,C., Serrano,R., Soler,G., Tadeo,F., Talon,M.,
            Terol,J., Trenor,M., Vaello,L., Vicente,O., Vidal,Ch., Zacarias,L.
            and Conejero,V.
  TITLE     Development of a citrus genome-wide EST collection and cDNA
            microarray as resources for genomic studies
  JOURNAL   Plant Mol. Biol. 57 (3), 375-391 (2005)
  PUBMED    15830128
COMMENT     Contact: Forment J
            Genomics Laboratory
            Instituto de Biologia Molecular y Celular de Plantas (Universidad
            Politecnica de Valencia - Consejo Superior de Investigaciones
            Cientificas)
            Avenida de los Naranjos s/n, 46022 Valencia, Spain
            Email: jforment@ibmcp.upv.es.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Citrus sinensis"
                     /mol_type="mRNA"
                     /db_xref="taxon:2711"
                     /clone="C08018A08"
                     /sex="hermaphrodite"
                     /dev_stage="2-3 month old seedlings, 20-30 cm. high"
                     /lab_host="Escherichia coli"
                     /clone_lib="PhyRootSw1"
                     /note="Organ: roots; Vector: Uni-ZAP XR; cDNA library made
                     from a mixture of equal amounts of poly-A+ RNA from roots
                     of plant inoculated with Phytophthora citrophtora
                     zoospores by inmersion in a suspension with 2000 or 15000
                     zoospores/mL, extracted at 0, 2, 4, 8, 13, and 18 days
                     after inoculation"
ORIGIN

  Query Match             76.8%;  Score 19.2;  DB 8;  Length 680;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+03;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTAC 25
```

```
                     ||||| |||| |||||||| ||||
Db         66 AGACAATTCAAGAAGTTCACCTAC 89


RESULT 14
CK000827/c
LOCUS       CK000827                 770 bp    mRNA    linear   EST 26-NOV-2003
DEFINITION  AGENCOURT_16108238 NIH_MGC_221 Homo sapiens cDNA clone
            IMAGE:30708096 5', mRNA sequence.
ACCESSION   CK000827
VERSION     CK000827.1  GI:38526861
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin, University of Iowa
             cDNA Library Preparation: M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM1073  row: m  column: 01
            High quality sequence stop: 643.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30708096"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_221"
                     /note="Organ: mixed; Vector: pYX-Asc; Site_1: EcoRI;
                     Site_2: NotI; Library is oligo-dT primed and directionally
                     cloned Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according tomRNA size fraction, ligated with
                     EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. Average insert size
                     4-5Kb. Adaptors 5'(AATTCGGCACGAGG)3' and 5'd
                     (CCTCGTGCCG)3'. 3' Linker sequence - GCGGCCGCTGAGAGCC T18.
                     Sequencing primers 3'end: T3 promoter primer 5'd
                     (ATTAACCCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd
                     (TAATACGACTCACTATAGGG)3'. Library was constructed in the
                     laboratory of M. Bento Soares. Note: this is a NIH_MGC
                     Library"
ORIGIN

  Query Match             76.8%;  Score 19.2;  DB 7;  Length 770;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+03;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AGACAGTTCATGAAGTTCATCTAC 25
              |||||||||||||||  |||| ||
Db        704 AGACAGTTCATGAAGAGCATCAAC 681


RESULT 15
DN023192
LOCUS       DN023192                 892 bp    mRNA    linear   EST 10-FEB-2005
DEFINITION  JGI_CAAR3995.fwd NIH_XGC_tropLiv1 Xenopus tropicalis cDNA clone
            IMAGE:7736465 5', mRNA sequence.
ACCESSION   DN023192
VERSION     DN023192.1  GI:59202606
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
            Brokstein,P. and Lindquist,E.A.
  TITLE     DOE Joint Genome Institute Xenopus tropicalis EST project
  JOURNAL   Unpublished (2004)
COMMENT     Other_ESTs: JGI_CAAR3995.rev
            Contact: Lindquist,E.A., Richardson,P.
            DOE Joint Genome Institute
            2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            Tel: 925 296 5600
            Fax: 925 296 5710
            Email: cdna@jgi-psf.org
            Tissue Procurement: Robert M. Grainger
            cDNA Library Preparation: Bruce Blumberg Laboratory, University of
            California, Irvine
            DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
            Clone Distribution: I.M.A.G.E. Consortium/LLNL:
            http://image.llnl.gov
            Naming Conventions: EST name is generated by the concatenation of
            the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
            indicates a forward sequencing read of the insert. It does not
            necessarily reflect the orientation of the insert.
            Plate: CAAR 0041  row: f  column: 15
            High quality sequence stop: 830.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /strain="N6 (Nigerian 6th generation inbred)"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:7736465"
                     /sex="male"
                     /tissue_type="Liver"
                     /dev_stage="Adult"
                     /lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
                     /clone_lib="NIH_XGC_tropLiv1"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
                     library was prepared from 5 ug of poly A+ RNA by oligo-dT
                     priming
                     (5'-GAGAGAGAGAGAGAGAGACTAGTCTCGAGTTTTTTTTTTTTTTTT-3')
                     and Stratascript reverse transcriptase. After ligation of
                     EcoRI adapters (5'-AATTCGGCACGAGG-3') followed by kinasing
                     adapters and by XhoI digestion, the cDNA was size selected
                     by chromatography on Sepharose CL-2B columns and fractions
                     containing cDNAs larger than 1000 bp were ligated into
                     EcoRI/XhoI-digested pCS107. Reference for library
                     construction: Current Genomics 4, 635-644.  Library
                     constructed by Michelle Tabb and Bruce Blumberg (Dept of
                     Developmental and Cell Biology, University of California,
                     Irvine)."
ORIGIN

  Query Match             76.8%;  Score 19.2;  DB 8;  Length 892;
  Best Local Similarity   84.0%;  Pred. No. 1.1e+03;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 TAGACAGTTCATGAAGTTCATCTAC 25
              |||| ||||||  |||||| |||||
Db        831 TAGATAGTTCANAAAGTTCTTCTAC 855


Search completed: February  6, 2006, 15:50:28
Job time : 2332 secs
```